Translating Big Data into Better Information for TEM: The Camera Evolves

Cory Czarnik
Gatan Inc.

As we transition from the previous generation of cameras and detectors (e.g., CCD based sensors, relatively slow frame rates, binned images to improve frame rate) to the current generation of cameras (large format CMOS detectors, video-like frame rates and extensive in-situ imaging), the demands on data workflow and data management are growing exponentially. Historically, it had only been practical to record a "picture" of an event in the TEM (e.g., VHS tape, screen capture software, etc.) whereas today's camera systems are capable of storing the data in each (of perhaps 16M) pixel(s) at tens of frames per second directly to disk for offline quantitative analysis. This is facilitated by the rapid increase in computational power, data transfer efficiency, and data storage capabilities at the same time that the cost of these capabilities approaches something that can be incorporated into a single (albeit high end) PC with a commercial operating system. This has opened up many new applications including identifying fine detail in structural biology specimens (including the recent Zika virus), 4D STEM application to strain mapping in materials, sub-ms temporal resolution for in-situ reaction studies, and analysis of beam-sensitive zeolite and metal-organic framework samples as well as simply improving the usability of cameras for "picture taking" with fast frame drift correction of images.

It is increasingly clear that a streamlined, robust, yet simple method of managing and analyzing large data sets needs to be developed in parallel with the applications themselves. For example, the structural biology community is currently employing direct detection cameras that output 80Gb/sec from the camera that is subsequently processed and reduced in size in order to make it useful for quick-turn analysis. While it may be straightforward to use a "brute force" method of storing, moving, and analyzing a single data set to demonstrate proof of concept for a given application, optimizing the workflow for each application will be required for widespread adoption and implementation.